NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

SUBMARINE BASE, GROTON, CONN.

REPORT NO. 1046

DETECTION AND RECOGNITION PERFORMANCE OF SONAR OPERATORS IN A MULTIMODAL TASK

by

D. A. Kobus, J. Russotti, C. Schlichting,
G. Haskell, S. Carpenter,
and
J. Wojtowicz

Naval Medical Research and Development Command
Research Work Unit M00~~96.002-1047~~
100.011-1021

Released by:

W. C. Milroy, CAPT, MC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory

26 April 1985

DETECTION AND RECOGNITION PERFORMANCE OF SONAR OPERATORS IN A MULTIMODAL TASK

by

D. A. Kobus, LT, MSC, USNR
J. Russotti, M.A.
C. Schlichting, Ph.D.
G. Haskell, Ph.D.
S. Carpenter, B.A.
J. Wojtowicz

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY
REPORT NUMBER 1046

NAVAL MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Research Project M0100.001-1021

Approved and Released by

W. C. MILROY, CAPT, MC, USN
Commanding Officer
NAVSUBMEDRSCHLAB

SUMMARY PAGE

PROBLEM

To determine if auditory, visual or the multimodal approach is best for detection and classification of "real world" targets. Actual auditory and visual sonar displays have not been used in previous investigations and, therefore, were used in the present study.

FINDINGS

The results indicated that the best modality for detection was target specific. However, detection performance in the multimodal condition was not significantly different from the best single modality for a given target.

APPLICATION

The finding that the best modality for detection was target specific, and that the multimodal approach was not significantly inferior to the best single modality, lead to the conclusion that the multimodal approach is best for initial target detection in the operational setting.

ADMINISTRATIVE INFORMATION

This research was conducted as part of the Naval Medical Research and Development Command Work Unit M0100.001-1021- "Auditory Sonar." It was submitted for review on 12 Mar 1985, approved for publication on 26 April 1985, and designated as NSMRL Report No. 1046.

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

ABSTRACT

Trained sonar operators participated in a detection and classification task. Stimuli were presented in three conditions: auditory and visual modalities independently and simultaneously (multimodal). Elapsed time and signal-to-noise (S/N) ratios were recorded. The best modality for target detection was found to be target specific. However, the multimodal condition was not significantly different from the single best modality and, therefore, should be used for initial target detection in the operational setting. The difference from findings in previous studies is discussed.

INTRODUCTION

Sonar has historically used the auditory modality to present acoustic data to an operator. Over the past decade, however, visual displays have also been developed. Sonarmen must now interpret complex auditory and visual information which is presented simultaneously. This information may or may not be meaningful (i.e., related to a target of interest). Yet most research that has been done to enhance sonar performance has investigated only a single modality. An important question, which has been virtually overlooked, is how two types of information (such as aural and visual) are processed when presented simultaneously.

The few studies which have involved two or more sensory modes have used very simple stimuli (1), and only a handful have come close to a method of presenting more than one mode of meaningful information simultaneously (2,3). Moreover, these studies tend to use a method of directed attention, such that subjects attend to one stimulus mode (visual) while an incidental mode (aural) of stimulation competes (4).

These experiments have provided a useful foundation for multimodal research, but they have not addressed two of the more salient issues: (a) What are the effects of simultaneous presentation of two or more modes of stimulation? and (b) How is performance affected when the stimuli are meaningful (i.e., sonar signals, numbers, words, colors, and symbols)?

A number of studies has shown that reaction time is faster with multimodal stimulation (5-8), and other studies have found that signal detection is improved when the information is presented multimodally (9-12). Indeed, in no case has performance with multimodal presentation been found to be significantly inferior to that with a single mode display (13,14). Hanson (4) has concluded that there are specific codes used in both the visual and auditory modalities, and that information received in one modality definitely has an enhancing effect by reducing reaction times on the other.

Only within the last decade has a study investigated the applied situation in a more detailed manner. Colquhoun (14) evaluated long duration performance using actual sonar sounds to provide an aspect of realism. This technique had not been used in previous studies, and may have influenced the results. Yet, the visual display he used, a simulated picture of "vertical tracks" on a screen, was not completely realistic. His results indicated that in the vigilance situation, overall detection performance was <u>best</u> when the information was presented simultaneously to both modalities. Thus, it appears that information from the two modalities may combine in some way to reduce detection threshold.

Another possibility should be kept in mind. Jaquish (15) has proposed that people are differentially attuned to the sensory worlds of

sound, sight, and touch, and different individuals respond best to stimuli they are most proficient with. For example, a photographer would respond best to visual stimulation. Perhaps a sonar operator may respond best to multimodal stimulation.

The primary goal of this study was to determine if the auditory, visual, or the multimodal approach is best for detection and classification of "real world" targets. Actual auditory and visual sonar displays were employed to investigate detection and classification performance of trained sonarmen.

METHOD

Subjects: Nine highly trained sonar operators volunteered to serve as subjects. All had or were corrected to 20/20 visual acuity and displayed hearing within the normal range in routine audiometric testing.

Apparatus: Testing was conducted using a multi-channel target simulator and operator console. The simulator contained a microprocessor which controlled signal intensity of recorded targets and background noise levels. The signal processing simulated that of a sonar system.

The target signal was provided by a Scully 284B-8 tape transport which was fed to the sonar simulator. A Scully 280B-2 tape transport supplied the recorded background signal. All target signals were continuous recording loops of specific sonar targets. The targets were generated by rule, using signal generation techniques available at the U.S. Navy Sonar Operational Trainer (SOT) at the Naval Submarine School. Target signals were recorded with an accuracy of $\pm$ .5 dB across the spectrum. Background sea-noise was also recorded using the same specifications.

The nominal broadband signal-to-noise ratio (S/N) of each target, when the operator was trained on it, was 0 dB S/N referred to the background noise. Digital attenuation of each target channel, up to 40 dB, reduced the maximum signal level. The simulator was programmed to increase target S/N, thereby simulating a closing target at a selected rate (dB/min.). A handwheel and bearing indicator allowed the operator to train on and off the target. All signals were well below threshold at the beginning of each trial. The simulator output simultaneously provided both an auditory signal on a Koss PRO-4-AAA headset, and a visual signal on the AN/BQR-20A. The latter generated a time-history display of the frequency spectrum of the sonar signal. Frequency is displayed along the x-axis with time along the y-axis. New information was displayed every 500 ms as a single raster line of data across the top of the display. The data moved in a "waterfall" fashion over time, taking 16 s to update the complete display. Frequencies detected by the system appeared as lighted dots along the x dimension, with amplitude coded by the intensity of the dot. A vertical cursor,

controlled by the subject, allowed numeric readout of specific frequency information. Figure 1 is a photograph of a typical display.

Procedure: A training session was followed by three experimental conditions: auditory only, visual only, or simultaneous (multimodal) exposure. The order of conditions was counterbalanced over subjects.

Subjects read a description of the task prior to training. During the training session, they listened to and viewed all targets. They were given as much time as necessary to become confident that they could recognize the targets. Subjects were then told that they must determine the target bearing within one degree to be scored correct for auditory detection, or within 5 Hz for a correct visual detection.

Testing: A two-way communication system, with an open microphone located in the testing room, effectively isolated the subject. Target location was randomized and, prior to the start of each trial, the operator was given a 10 degree sector to search. This was done by moving the handwheel which allowed the operator to select one degree at a time. All targets were presented individually and were initiated at a -20 dB S/N which increased at a rate of 3 dB/min. The same five targets were presented to each subject, in different random orders, in each of the three (auditory, visual, multimodal) conditions. The experimenter cued the subject when each trial began. Subjects were instructed to verbally report detection and classification as soon as possible. Time and S/N when the targets were correctly detected were recorded, and the subject was then directed to continue observing the target until he provided the correct classification. Similarly, time and S/N were recorded for correct classification. All incorrect responses resulted in a "negative, please continue" instruction from the experimenter. This procedure continued until all trials were completed within a session. A 10-minute rest period was provided between conditions.

RESULTS

Detection

Detection performance in the multimodal condition was not significantly different from the best single modality for a given target.

The mean S/N ratio for detection for the auditory condition was -12.17 dB, for the visual condition -9.88 dB, and -12.13 dB when the information was presented multimodally. These differences were not statistically significant ($F(2,16)=2.95$, $p<.10$). S/N ratio was significantly different among the five targets for detection performance ($F(4,32)=6.32$, $p<.01$). Additionally, the S/N level at which a target was detected significantly interacted with the mode of stimulus presentation ($F(8,64)=9.61$, $p<.01$). That is, one target was most

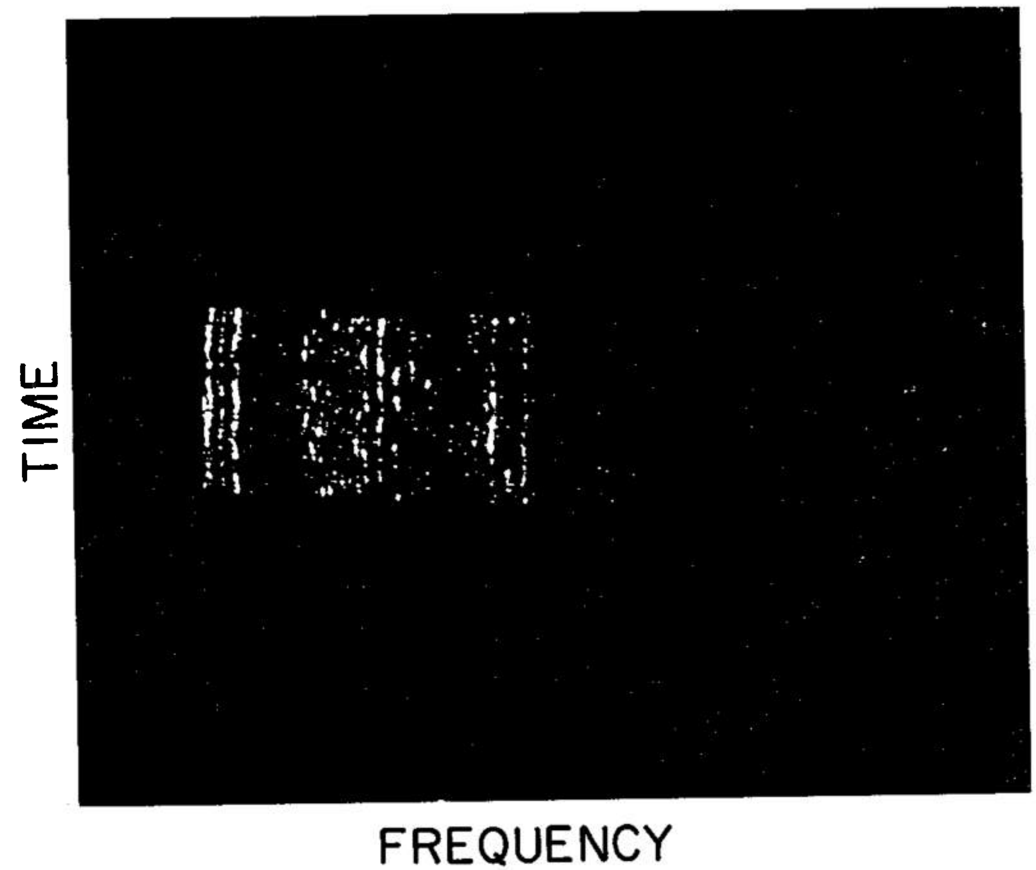

Figure 1. Typical visual presentation from one of the targets. The signal shown is well above threshold. Upon initial detection a target may consist of only one "vertical signal." Time is displayed along the y dimension with frequency along the x dimension.

difficult to detect in the visual modality (#3), while another target (#4) was most difficult to detect in the auditory modality. These results are displayed in Figure 2.

Figure 3 shows the results of the mean elapsed time required to detect the target. This measure was used to provide a more direct comparison with previous research. Since S/N ratio was directly related to elapsed time in our paradigm then similar trends were observed. More time was taken to detect targets in the visual condition (207.6 s) than in the other two conditions (auditory, 157.0 s; multimodal, 159.5 s). These differences were not statistically significant ($\underline{F}(2,16)=3.43$, $\underline{p} < .10$). As with the S/N data, there were significant differences among targets ($\underline{F}(4,32)=5.73$, $\underline{p} < .01$) and the interaction of target and condition was also significant for this analysis ($\underline{F}(8,64)=10.39$, $\underline{p} < .01$).

A Newman-Keuls analysis was performed on the S/N ratio data to investigate individual target differences. Visual detection and classification performance was significantly poorer ($\underline{p} < .01$) for target #3 than any other target. Visual performance for target #1 was significantly better ($\underline{p} < .05$) than all targets except #4 for detection and targets #2 and #4 for classification. Auditory detection and classification performance, on the other hand, was significantly better for target #3 ($\underline{p} < .05$) than the other targets with the exception of target #2. No significant differences between targets were found for detection and classification performance using the multimodal condition.

Classification

Figure 2 also shows the mean S/N ratio at which the subjects were able to correctly classify the targets for each condition (shown by -C-). In agreement with the detection data, the S/N ratio at classification, also displayed a significant difference between individual targets ($\underline{F}(4,32)=3.41$, $\underline{p} < .05$). Although when the data were collapsed across targets no significant differences in S/N ratios between the three conditions were found. When individual target data was added to the analysis there was a significant (condition x target) interaction ($\underline{F}(8,64)=8.93$, $\underline{p} < .01$).

Elapsed time to correctly classify a target is marked in Figure 3. Mean time until the target was correctly classified took the longest in the visual condition (224.6 s) and took the shortest amount of time in the auditory condition (189.6 s). The time to classify in the multimodal condition (212.7 s) was longer than in the auditory condition alone but shorter than in the visual condition. However, the differences between these values were not statistically significant.

On the average subjects took approximately 32 s to classify a target in the auditory condition after they had detected it. In the visual condition this time was much shorter, 17 s. In the multimodal

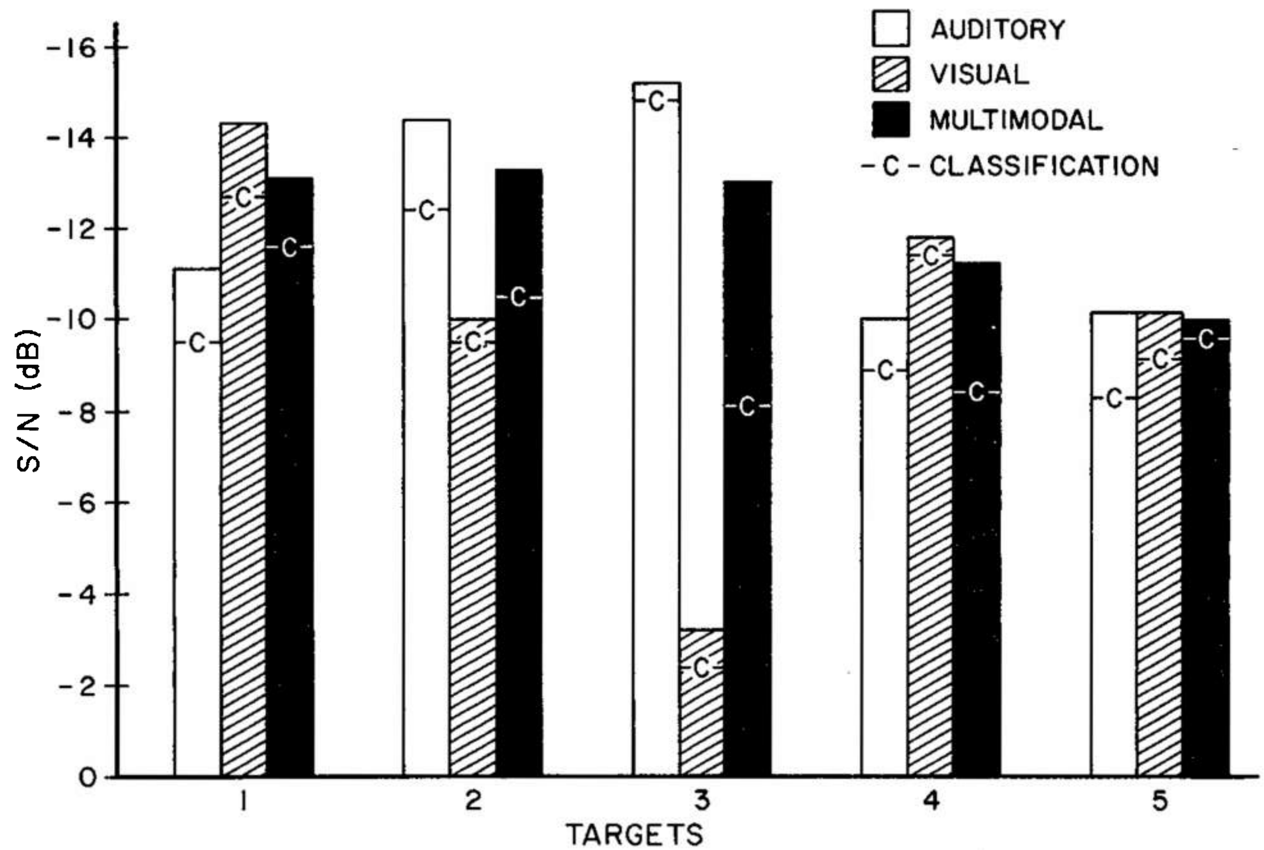

Figure 2. Mean S/N values for each target and for each condition when the targets were initially detected. Symbol -C- indicates the value when the targets were classified. (Open bar = auditory; crossed bar = visual; filled bar = multimodal).

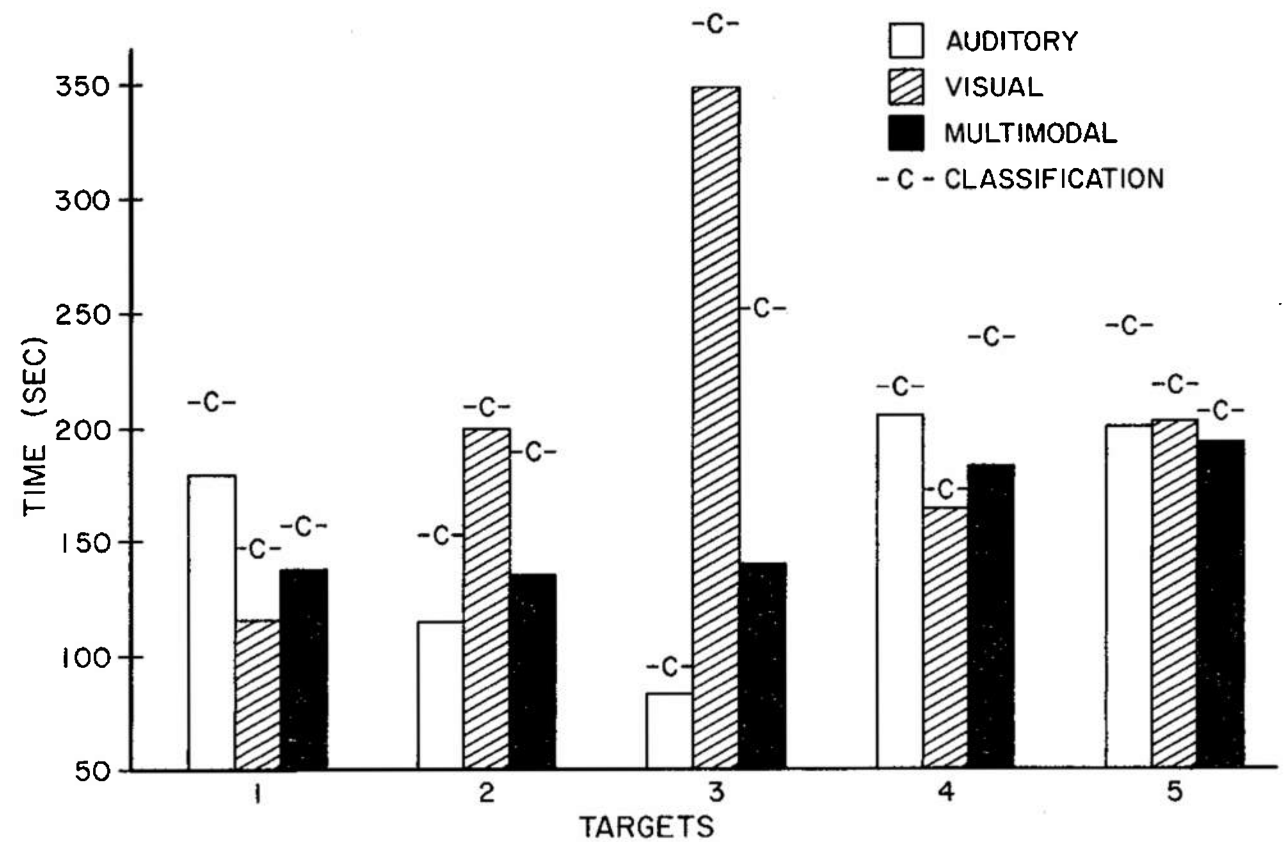

Figure 3. Mean elapsed time for target detection and classification for each condition. Classification times are indicated by the symbol -C-.(open bar = auditory; crossed bar = visual; filled bar = multimodal).

condition, however, the subjects averaged more than 53 s to classify after detecting the target. When the data were separated for individual target analyses the time to classify each target was significantly different between the conditions resulting in a significant (target x condition) interaction ($\underline{F}(8,64)=10.35$, $\underline{p} < .01$).

Particular targets were detected and classified faster in specific modalities. This result is similar to the detection interaction discussed previously. The relationship was fairly consistent showing that the order in which the targets were classified was similar to the order in which they were detected.

DISCUSSION

This study was the first to investigate multimodal detection and classification performance using the actual methods of presenting sonar information that are used in the applied setting. Our results agree with previous claims that multimodal presentation is not significantly different from the better single modality's result. In addition, the present results have shown that the better modality for detection or classification is target dependent. In other words, the auditory modality was better for some targets, but the visual modality was better for others. This result is probably due to the large differences in spectral characteristics between the targets. However, the stimuli used in this study were chosen to provide a representative sample of actual targets that an operator may be exposed to during a routine watchstanding period. More importantly, individual target differences were not found using the multimodal condition. Therefore, the finding that the best modality for detection was target specific, and that the multimodal approach was not significantly inferior to the best single modality, leads to the conclusion that the multimodal approach is best for initial target detection in the operational setting. This finding is of additional importance due to the recent de-emphasis placed upon auditory sonar detection.

Although there were large differences in the time required to classify a target between the experimental conditions, it should be pointed out that this relationship was also target specific. In most cases, if an advantage was shown for one modality, then the target was both detected and classified prior to being detected by the other modality. However, this relationship did not hold true for the multimodal condition. Regardless of which modality provided the faster detection times, the multimodal condition was not significantly different. This finding also supports using the multimodal approach for target classification.

The present results failed to support the conclusion of Colquhoun (14) regarding an enhancement in detection performance when the information was presented multimodally. His findings may have resulted from using a simulated rather than an actual visual display. Also, his subjects were selected from "various Navy categories" and may

have had different processing strategies than those used by the highly trained sonarmen in this study.

A problem with the present study was the lack of control over exactly when a target's information would exceed threshold in either modality. In addition, the signals were extremely complex. These aspects were necessary to provide an attempt at realism. However, in order to provide a better understanding of multimodal processing a more progressive approach should be applied. First, an additional experiment should be conducted in which the targets are simple, meaningful signals. The targets should be controlled in such a way that the onset (i.e., when the signal exceeds threshold) in each modality can be simultaneously or successively presented. Also, the signals should be made progressively more complex. Such a study might determine how the temporal order of stimulus input affects target detection and recognition performance.

REFERENCES

1. Nickerson, R. S. (1983). Intersensory facilitation of reaction time: energy summation or preparation enhancement? Psychological Review, 80, 489-509.
2. Keele, S. W. and Neill, W. T. (1978). Mechanisms of attention. In E. C. Carterette & M. P. Friedman (Eds.) Handbook of Perception: Vol.IX. Perceptual processing (pp 3-7). New York: Academic Press.
3. Burrows, D. and Solomon, B. A. (1975). Parallel scanning of auditory and visual information. Memory and Cognition, 3, 416-420.
4. Hanson, V. L. (1981). Processing of written and spoken words: evidence for common coding. Memory and Cognition, 9, 93-100.
5. Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. Journal of Experimental Psychology, 63, 289-293.
6. Helson, H. (1964). Current trends and issues in adaptation-level theory. American Psychologist, 19, 26-38.
7. Morrell, L. K. (1967). Intersensory facilitation of reaction time. Psychonomic Science, 8, 77-78.
8. Bernstein, I. H., Rose, R. and Ashe, V. M. (1970). Energy integration in intersensory facilitation. Journal of Experimental Psychology, 86, 196-203.
9. Brown, A. E. and Hopkins, H. K. (1967). Interaction of the auditory and visual sensory modalities. Journal of the Acoustical Society of America, 41, 1-6.
10. Fidell, S. (1970). Sensory function in multimodal signal detection. Journal of the Acoustical Society of America, 47, 1007-1015.
11. Loveless, N. E., Brebner, J., and Hamilton, P. (1970). Bisensory presentation of information. Psychological Bulletin, 73, 161-199.
12. Buckner, D. N. and McGrath, J. J. (1963). A comparison of performance on single and dual sensory mode vigilance tasks. In D. N. Buckner and J. J. McGrath (Eds) Vigilance: A symposium. Pp.53-69. New York: McGraw-Hill.

13. Baker, R. A., Ware, J. R., and Sipowicz, R. R. (1962). Vigilance: a comparison in auditory, visual, and combined audio-visual tasks. Canadian Journal of Psychology, 16, 192-198.
14. Calquhoun, W. P. (1975). Evaluation of auditory, visual, and dual-mode displays for prolonged sonar monitoring in repeated sessions. Human Factors, 17, 425-437.
15. Jaquish, G. A. (1983). Intra-individual variability in divergent thinking in response to audio, visual, and tactile stimuli. British Journal of Psychology, 74, 467-472.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NSMRL Report No. 1046	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DETECTION AND RECOGNITION PERFORMANCE OF SONAR OPERATORS IN A MULTIMODAL TASK		5. TYPE OF REPORT & PERIOD COVERED Intermin Report
		6. PERFORMING ORG. REPORT NUMBER NSMRL Rep. No. 1046
7. AUTHOR(s) D. A. Kobus, J. Russotti, C. Schlichting, G. Haskell, S. Carpenter, J. Wojtowicz		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Submarine Medical Research Laboratory Naval Submarine Base New London Groton, Connecticut 06349-5900		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 65856N M0100.001-1021
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Submarine Medical Research Laboratory Naval Submarine Base New London Groton, Connecticut 06349-5900		12. REPORT DATE 26 April 1985
		13. NUMBER OF PAGES 10
14. MONITORING AGENCY NAME & ADDRESS (If different from Controlling Office) Naval Medical Research and Development Command Naval Medical Command, National Capital Region Bethesda, Maryland 20814		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multimodal; sonar performance; detection; recognition

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Trained sonar operators participated in a detection and classification task. Stimuli were presented in three conditions: auditory and visual modalities independently and simultaneously (multimodal). Elapsed time and signal-to-noise (S/N) ratios were recorded. The best modality for target detection was found to be target specific. However, the multimodal condition was not significantly different from the single modality and, therefore, should be used for initial target detection in the operating setting. The

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. continued

difference from findings in previous studies is discussed.